REMOTE SENSING CENTRE

1983 MANITOBA CROP AREA ESTIMATION PROJECT

L. Horn
R. A. Ryerson
C. Thibault
R. Dobbins

Manitoba
Natural Resources
Surveys and Mapping

May, 1984

1983 MANITOBA CROP

AREA ESTIMATION PROJECT

L. Horn

R.A. Ryerson

C. Thibault

R. Dobbins

A cooperative project under the
Technology Enhancement Program
between Manitoba Remote Sensing
Centre, Canada Centre for Remote
Sensing, Statistics Canada and
the Manitoba Department of Agriculture

May, 1984

ABSTRACT

Under the terms of the Technology Enhancement Program (TEP), the Canada Centre for Remote Sensing (CCRS) and the Manitoba Remote Sensing Centre (MRSC) have conducted a cooperative crop monitoring project with the Manitoba Department of Agriculture and Statistics Canada. The purpose of the project was to assess the feasibility of using remote sensing techniques for crop reporting purposes in Manitoba. The specific objectives of the project were to produce timely area estimates for canola and to conduct a test/demonstration of grain and fallow area estimation in Crop Districts 1, 2 and 3.

Information on crop types and their exact location provided by an in-house Statistics Canada probability enumerative survey served as the basis for interpreting satellite data and assessing classification results. All image analysis was performed with the Applied Resource Image Exploitation System (ARIES-II) and integration of classification results with survey information was accomplished by linear regression.

Results of the project showed that remote sensing techniques could be used to produce timely crop area estimates at the sub provincial level which were comparable to the National Farm Survey and Crop Reporting estimates. Generally, the low remote sensing estimate for rapeseed was attributed to differences between image dates and crop conditions and to training area purification and subsequent reduction in spectral signature. Higher remote sensing estimates for summerfallow and cereals were attributed to inclusion of confusion cover types. In the summerfallow category, roads and sparsely vegetated areas were included, whereas in the cereal category, some areas of natural grasslands, improved pasture, flax and corn were included.

ACKNOWLEDGEMENTS

This project was made possible through the initiatives of W.G. Best, Chief of the Manitoba Remote Sensing Centre and Coordinator for the Manitoba Technology Enhancement Program, J. Douglas Heyland, Chief of Technology Enhancement, CCRS, Barry Proud, Chief (now on leave) of the Crops Section, Statistics Canada and E.T. Lewis, Chief, Statistical Analysis Section, Manitoba Agriculture. Technical inputs and advice were received from Messrs. M. Shumsky, O. Code, R. Plourde and B. Snider of Statistics Canada and from Christian Prévost of CCRS.

Appreciation is also extended to the staff at the Prince Albert Satellite Station (PASS) for their assistance and cooperation in acquisition of timely Landsat data and to the NFS enumerators for their good work in gathering ground reference data. In addition, many thanks to the staff of the Cartography Section of the Manitoba Surveys and Mapping Branch for their assistance in preparing figures and maps for the report.

A special note of thanks is extended to L. Kiselbach of the Manitoba Surveys and Mapping Branch for typing this report.

A NOTE ON AUTHORSHIP

Although performed by the authors, responsibility for publication and distribution of this report rests with the Manitoba Remote Sensing Centre. L. Horn of the Canada Centre for Remote Sensing was on secondment to the Manitoba Remote Sensing Centre for the duration of the project. R.A. Ryerson also with the Canada Centre for Remote Sensing was on secondment to Statistics Canada when this work was being conducted. Both C. Thibault and R. Dobbins are with Statistics Canada.

CONTENTS

	Page
ABSTRACT	*i*
ACKNOWLEDGEMENTS	*iii*
LIST OF TABLES	*v*
LIST OF FIGURES	*vi*
1. INTRODUCTION	1
1.1 Background	1
1.2 Objectives	1
1.3 Approach	2
1.4 Study Area	3
2. METHODS	3
2.1 The National Farm Survey (NFS) Design	3
2.1.1 NFS Target Populations and Sampling Frame	5
2.1.2 Area Sample Design	6
2.2 Acquisition of Ground Reference Data	7
2.3 Acquisition of Satellite Imagery	9
2.4 Image Analysis Procedures	10
2.4.1 Preparation for Analysis	10
2.4.2 Signature Development and Classification	14
2.4.3 Post-Classification Analysis	19
2.5 Crop Area Estimation	25
2.5.1 Post Stratification	25
2.5.2 Regression Estimation	26
3. RESULTS	28
4. DISCUSSION	33
4.1 Area Estimation	33
4.2 Sample Design	37
4.3 Post Classification Filtering	38
4.4 Boundary Placement	40
4.5 Feasibility of a Provincial Crop Monitoring Program	42
5. CONCLUSIONS	45
6. RECOMMENDATIONS	47
LITERATURE CITED	48
APPENDIX I - Interviewers Instruction Manual.	
APPENDIX II - Thematic Map of Image TC1 Showing Filtered and Unfiltered Results.	
APPENDIX III - Pixel Count by Segment for Both Filtered and Unfiltered Results for Images SWQ, TC1 and RIM.	
APPENDIX IV - Detailed Account of Pixel Count by Segment for Image SWQ.	

LIST OF TABLES

Table		Page
1.	Codes for crop types and land cover identified on air photographs by NFS enumerators.	8
2.	Remote sensing regression estimates for SPA1 and each of the satellite images.	29
3.	1983 Manitoba Crop Monitoring Remote Sensing estimates for Crop Districts 1, 2 and 3.	31
4.	1983 Manitoba Crop Monitoring study survey estimates.	34
5.	Hierarchy of soil factors which affect wheat yields.	39

LIST OF FIGURES

Figure		Page
1.	Study area for the 1983 Crop Monitoring Project.	4
2.	Scattergram showing the pattern and location of spectral clusters associated with six training sets for image TC1.	17

1. INTRODUCTION

1.1 Background

The 1983 Manitoba Crop Monitoring study was developed as a cooperative project under the joint Federal/Provincial Technology Enhancement Program (TEP). The objective of the program was to enhance the scientific and technical capabilities of remote sensing in Manitoba and to apply this technology to specific resource management issues. Implicit within the structure of the TEP program was that CCRS as technology developers would conduct demonstration projects with potential user agencies who would then assume the role of technology adopters. The first year of each project in the program would be carried out as a joint pilot study and in subsequent years the project would expand to a larger scale up to the point where user agencies would adopt the technology as part of their operational program.

The purpose of this report is to document the techniques and results of the 1983 pilot project for crop monitoring in Manitoba. Participants in the project were the Canada Centre for Remote Sensing (CCRS), Manitoba Department of Agriculture (MDA), Manitoba Remote Sensing Centre (MRSC) and Statistics Canada.

1.2 Objectives

The purpose of the study was to demonstrate the feasibility of using remote sensing techniques for crop reporting purposes. The specific objectives of the project were:

1. To produce timely area estimates for canola in Crop Districts 1, 2 and 3.

2. To conduct a test/demonstration of grain and fallow area estimation in Crop Districts 1, 2 and 3.

1.3 Approach

Information on crop types and their exact location was provided by an in-house Statistics Canada probability enumerative survey. This data served as the basis for interpreting satellite data and assessing the classification results. Area calculations for each target crop were made using a regression estimation technique (Hanuschak et al 1980). The details of the area estimation procedure have been described elsewhere (Ryerson et al 1981, Ryerson et al 1983).

The National Farm Survey (NFS) is a multipurpose survey conducted each year in Canada that is designed to produce reliable provincial estimates for a variety of crops, livestock and farm operating expenses. Details of the methods used in designing the survey are given by Ingram and Davidson (1983). In addition to the standard questionnaire, NFS enumerators were instructed to identify and record all crops and their acreage onto enlarged aerial photographs for each segment.

Analysis of satellite data was performed with the ARIES-II image analysis system. A supervised classification technique which uses the Maximum Likelihood Decision Rule (MLDR) to make class discriminations was used to interpret the satellite data (Goodenough and Shlien 1974).

Satellite based crop area estimates were compared to the segment crop areas and a linear regression was used to adjust the raw satellite estimate. This method has proven to be more accurate for generating area estimates than either satellite or survey data alone (Hanuschak et al 1980).

1.4 Study Area

The study area was located in the southwest corner of the province and represented a land area of approximately 24,000 km² (Figure 1). Over 1,809,000 hectares representing 75 percent of the land area has been developed for agricultural purposes (Manitoba Agriculture 1981). The study area included all of Manitoba Crop Districts 1, 2 and 3 and was considered for NFS purposes as one sub-provincial area (SPA).

2.0 METHODS

The integration of the NFS survey with remote sensing techniques to produce crop area estimates was accomplished in five major steps. The following sections highlight the key elements of each step. In particular, the NFS survey design will be highlighted, the methods used to acquire ground reference and satellite imagery will be noted, image analysis procedures will be documented and finally, the area estimation technique will be shown.

2.1 The National Farm Survey (NFS) Design

The objective of the NFS is to provide provincial estimates of level and change for crop, livestock and farm operating expense items. In addition, the NFS serves as an annual benchmark and sampling frame for integrating other surveys such as crop commodity and other livestock surveys.

It is important to note that the NFS is a multi-purpose survey designed to satisfy a number of information requirements for agricultural purposes. The fact that the survey was designed to meet agricultural information needs does not mean, however, that it cannot be used to assist in

Figure 1. **Study area for the 1983 Manitoba Crop Monitoring project. The numbers 32 and 33 refer to Landsat orbital path numbers.**

interpreting satellite data. Examination of the sampling frames, sample design and stratification procedures of the NFS as well as special considerations given to remote sensing needs provides further insight into the integration of remote sensing with the NFS.

2.1.1 NFS Target Populations and Sampling Frames

In Manitoba, the NFS target population was defined as all agricultural holdings which received $2,000 or more from the sale of agricultural products during the previous 12 months excluding:

a) farms located on Indian Reserves,

b) institutional farms, and

c) farms in marginal areas with little or no agricultural activity (determined from the 1981 census data).

The multiple frame sampling approach adopted for the NFS consisted of samples selected from two list frames and one area frame (Ingram and Davidson 1983).

The core survey, made up of an area and list frame, was built at the provincial level and its purpose was to produce provincial estimates. In the initial list frame two selection criteria were of particular importance. First, each farm was to be large with respect to at least one of the following key variables: total cattle, total pigs, total cropland and total sales. Second, the list frame farms, collectively, accounted for a minimum percentage of the provincial target population for each key item.

The extended list frame or Mail Telephone Follow-up (MTF) survey was built at the sub-provincial area level using the same key variables. The goal of the MTF list frame design was to account for approximately the same minimum percentage of the target population total at the SPA level as had

been accounted for at the provincial level by the core list frame.

The area frame consisted of all the 1981 Census Agriculture Enumerative Areas (EA's) in excluded areas. An EA is the geographical area canvassed by an enumerator during the census.

The area frame was used in conjunction with either the list or extended list frame for generating NFS estimates. For remote sensing purposes only the area frame was required to generate the estimates.

2.1.2 Area Sample Design

The area sample designed at the SPA level was implemented independently from the core list and MTF sample designs in each SPA. The design involved the selection of a stratified, replicated, two stage sample of parcels of land (segments) with 1981 census EA's as first stage sampling units and segments within EA's as second stage sampling units.

In the first stage design, farm level data from the 1981 Census of Agriculture were summarized to the EA level and were used to stratify the EA's and to allocate the sample size to the resulting strata. Only farms that did not belong to the core list frame were used to create the EA summaries. A hierarchical cluster analysis using total cattle and total cropland as key variables was used to stratify the EA's.

Allocation of the area sample size to the SPA was accomplished in two steps. Initially, the optimum allocations were determined separately for each key item and then a suitable compromise to the optimum allocation was determined with the method developed by Chatterjee (1967). This method assumed that all key items were of equal importance, therefore, in cases where it was desirable to attach more importance to some key items than to others, the Chatterjee allocation was modified. Stratum sample sizes were adjusted to multiples of two which permitted the selection of two independant replicates of EA's per stratum.

The EA's comprising the area frame for the SPA were sorted according to stratum, then by crop district and finally by the cumulative sales figures for farms located within the EA. Then, within each stratum, a systematic sample of EA's with two independent replicates, with replacement, were selected.

In the second stage design, each selected EA was divided into a number of land areas called segments. Wherever possible, segment boundaries were marked by easily identifiable geographic features (i.e. roads, rivers, etc.) and contained segments with similar amounts of agricultural land as determined from topographical maps. The average segment size was three square miles. In Manitoba, one segment was selected without replacement for every 30 segments formed in each sampled EA.

In summary there were 191 EA's within the study area (SPA1). From this total, 76 different EA's were selected and among these 76, four were selected twice to arrive at a total sample size of 80. For each EA that was selected twice, two different sets of segments were selected. Within these EA's a total of 90 segments were selected.

2.2 Acquisition of Ground Reference Data

Ground reference data for the study was obtained by enumerators working within the sampled land segments of the National Farm Survey. Enlarged air photo mosaics of each 3-section segment were supplied to the enumerators.* While contacting the farmers for the NFS, enumerators were instructed to identify and record on these photographs the type of crops and the precise boundaries for all land holdings, three acres or larger, within each segment (Table 1). In addition, the acreages of each field or land holding as reported by the respondent were recorded. To ensure timely arrival of the "support" data in a consistant and reliable format, written

*Only a few segments were not 3 x 1 sections.

Table 1. Codes for crop types and land cover identified on air photographs by NFS enumerators.

OILSEEDS

F.....Flaxseed

M.....Mustard Seed

RA.....Rapeseed

SN.....Sunflower

GRAINS

B.....Barley

C.....Corn

O.....Oats

RY.....Rye

W.....Wheat

MG.....Mixed Grains (includes any mixture of wheat, oats, barley or rye)

GRASSES

A.....Alfalfa

CL.....Clover

G.....Grass/Pasture

OTHERS

S.....Summerfallow

OC.....Other Crops (i.e. peas potatoes, buckwheat, vegetables, etc.)

OL.....Other Land (i.e. barns, buildings, bushlands, creeks, headlands, roads, sloughs, etc.; applies to an area of a farm which is not used for agricultural purposes).

instructions which detailed precisely the procedures to be followed were included with an "Interviewers Instruction" manual (Appendix I).

2.3 Acquisition of Satellite Imagery

Two LANDSAT scenes from each of two satellite passes, indicated by the orbits overlaid on Figure 1 were required to give complete coverage of the area. Two scenes acquired on July 14 provided coverage of the west half of the study area and two scenes acquired on August 8 provided the remaining coverage. All four LANDSAT Computer Compatible Tapes (CCT's) had corrections applied for earth rotation and mirror velocity. Prior to the critical image acquisition period, arrangements were made with staff at the Prince Albert Satellite Station (PASS) to view the satellite data as it was received in real time. They would then report image quality to the MRSC on the day it was received. Images free of technical anomalies and with less than 20 percent cloud cover were considered to be acceptable. All images were delivered to the MRSC within 48 hours of the satellite overpass.

Previous work (Brown et al 1980) indicated that canola was usually separable from other crops during and after flowering. Hot and humid weather during the first two weeks of July promoted rapid crop development and by July 12 approximately 50 percent of the canola was flowering with the remainder nearing bloom stage. By July 19 canola was about 80-90 percent flowering. Two cloud free images covering the entire western region of the study area were obtained by PASS on July 14. Crop development reports released by the Economic Analysis Branch on July 26 stated that flowering was 90-100 percent complete for canola and that flowering was nearing completion by August 2. With less than 5 percent of the canola swathed by August 9,

the two LANDSAT scenes acquired by PASS on August 8 provided the remaining coverage of the study area. Swathing progressed to 20 percent by August 16 and to 60 percent by August 23.

2.4 Image Analysis Procedures

Computer analysis of satellite data was performed with the ARIES-II image analysis system using AIIASP software. Image analysis for crop Districts 1 and 2 was conducted at the Manitoba Remote Sensing Centre and analysis for Crop District 3 was conducted on the back-up system at the Canada Centre for Remote Sensing. Image analysis operations were conducted in three phases. Phase one consisted of five preparation steps. Phase two consisted of several tasks combined to result in a supervised classification. The third phase consisted of post classification filtering, area calculation of thematic data and generation of applicon plots.

2.4.1 Preparation for Analysis

The main objectives of the preparatory phase were to load the satellite data to disk, locate the line/pixel coordinates of the study area and subareas and create a permanent record on disk of the location of each ground reference segment. For each preparation step outlined below the rationale is highlighted, the required computer programs are outlined and the method of implementing the programs are described. In addition, attention is drawn to possible pitfalls and how to avoid them. The importance of backing up work after completion of each step cannot be overemphasized.

Step 1:

Load Images to Disk. Under program 'TL" load the image from CCT to disk and assign each new scene with a unique three character area name. The three character area names used in this project were SWA, TC1, RIM and CDM. The first letter of the three character area name should always be unique as it is this letter that is retained when forming subareas.

Examine the full scene under program 'ID', decimated if necessary, to ensure the data are free of anomalies. Inspect all four bands for problems of striping, dropped lines, pixel misregistration and ensure that each band has been assigned the correct four character identifier. The four character identifier most often used to describe feature files or raw bands was BND*. The * is "wild" and may be represented by any one of the band numbers. It is important to note that the time spent in developing a logical and well organized file naming convention will save time in the long run.

Step 2:

Locate Line/Pixel Coordinates of the Study Area. The purpose of this step was to record the line/pixel coordinates of the study area boundary or subareas (i.e. crop districts) on a base map. The base map then served as a handy reference for several subsequent steps. If working with a large image, it was necessary to use the subroutine 'MS' in program ID to load and reload, decimated and nondecimated images several times. When coordinates for the complete perimeter of the area had been determined, reload the image and delineate the complete study area. These boundaries will only be approximate for a decimated image but will be sufficient for identifying major discrepancies between line/pixel coordinates and the study area boundaries.

Step 3:

Create Homogenous Subareas. With the complete study area outlined of the video display, the next step was to stratify the region into homogenous units with the aid of soil, physiography or land use maps. These stratifications were essential to ensure accurate signature extensions and ultimately good classifications. Major land units were delineated on the video display and subareas for each were created under the program 'AP'. Often, visual inspection of satellite data yielded the region's major patterns; however, ancillary data helped to isolate local areas where subtle differences in crop development occured due to soil texture or salinity.

Dominant soil orders and great groups as shown in the Manitoba Atlas (1983) were the basic criteria used to stratify the crop districts. Black soils of the Chernozemic order and Luvic Gleysols of the Gleysolic order dominated most of the study area. Local variations were introduced by the Dark Gray Chernozens, Humic Gleysols and Black Solonetz of the Solonetizic order and Orthic Regosols of the Regosolic order. Local differences in terrain, texture and topography also influenced the stratification process.

Feature files for each subarea were stored on separate back up tapes. Later, training and signature files as well as classified results were added to these tapes as they were created.

Step 4:

Locate Segment Data on Satellite Image. Under program 'IT' each ground segment was located and written to disk as a training area file. Each segment was designated by its own four character area identifier, SE**, where

the *'s corresponded to the segment number. The upper right and lower left line/pixel coordinates for each segment were also recorded as this information was extremely useful when selecting line/pixel coordinates of images that were loaded at a later date. The advantage of designating each segment as a training area was that they could be called up under program IT at any time thereby eliminating the need to painstakingly locate the segment every time information from it was required.

When creating segment training areas it was easier to locate the segments on a clear, two feature image. The addition of an extra band in program IT did not necessarily facilitate easier location of segment boundaries. Added image sharpness was obtained by assigning more bits to each feature per displayed pixel by loading only a two feature image (i.e. bands 5 and 7). Greater sharpness and image clarity enhanced the analyst's ability to locate segments.

All segment training areas were then backed up onto the appropriate tapes at the end of each image analysis session.

Step 5:

Create Training Areas for Total Area Measurement. Three large training areas created under program IT were required to cover crop districts 1, 2 and 3. These training area files were used in the calculation of total area of canola, summerfallow and cereal grains under the area calculation program 'CT'. Care was taken to ensure that land areas were not missed and that training areas between crop districts did not overlap. Using the update option of program IT it

was necessary to update the training area for crop district 1 a total of 12 times in order to get complete coverage. The training area for crop district 2 was updated 17 times. It was also during this step that training areas were created around cloud or haze affected satellite data. These areas were then removed from the total area calculation process.

At the completion of pre-image analysis work, numerous files had been written to disk and were stored on a series of back up tapes. This phase of image analysis required extensive operator interaction time and very little CPU time. Although the video display monitor was completely occupied during this phase, it was still possible to run some batch jobs in the background and even a few small real time tasks.

2.4.2 Signature Development and Classification

The main objectives of the image analysis phase were to select training areas most representative of each cover type, create signature files and generate a classified image. Ground reference data played a critical role during this phase as it provided the basis for interpreting satellite data. The program IT was used to define and purify training areas and the statistics generation program 'TS" was used to examine training area "purity". All image classifications were performed with the Fast Maximum Likelihood Classifier under program 'CB'. In other areas for some applications it may be possible to use the so-called parallelepiped classifier DP, which is faster. Each geographically distinct subarea created in the previous phase was analyzed separately.

Under program IT, a subscene of 508 lines by 509 pixels known to contain segment data was displayed on the monitor as a two feature image. Band 7 was displayed in red and band 5 was displayed in green. Although only two bands were displayed, statistical signatures were calculated from the digital values in all four bands. The operator had to specify all four bands using the SF sub-routine in program IT.

Once the subscene was displayed, ground segment locations previously stored were reloaded using the 'DT' subroutine of program IT. All segments in the subscene were loaded under one training area index number; however, before moving onto the next training area index number the "Continue" statement had to selected.

Ground segment photo mosaics prepared by NFS enumerators showing the current land cover types and their areas were used to select large fields for training. Several fields, all of the same cover type, were drawn under one training area index number; however, it was important to record the index number assigned to each cover type. It was the index number that was later specified for creation of temporary signature files and purification of training areas while in program IT.

The image analysis system recognized each cover type according to its spectral signature. The spectral signature for each training set was calculated from the digital values in all four spectral bands. The key to successful image classification was to define each cover type's precise spectral signature and remove those pixels from the training set whose digital values fell outside the spectral signature of the cover type. Although four bands were used to generate the signature, a scattergram of band 5 versus band 7 gave a good indication of the training area purity.

Because every pixel within a training area relayed significant spectral information about the cover type to the computer it was important to be selective and exact when choosing and defining training areas. Several training areas for the same cover type were selected from different segments located on the same image. In doing so, slight differences in spectral reflectance due to site-specific conditions were accounted for in the training set. A training set was comprised of all training areas created for the same cover type.

Although training areas were carefully selected it was still possible that they could contain pixels whose reflectance values were not truely representative of the target cover type. For example, border pixels included in the training area often possessed a spectral signature that differed from the "true" signature because of the influence that neighbouring cover types had on the border pixel's total reflectance. To detect the presence of these non-conforming pixels a scattergram of reflectance values in band 5 versus band 7 was examined under program TS. An example of a scattergram showing clusters for six training sets is shown in Figure 2.

The scattergram was a standard X-Y plot representing two-dimensional vector space. The pattern and location of each cluster were important for locating non-conforming pixels and for understanding where confusion between cover types might be expected to occur on classified images. Confusion could be expected whenever two spectral clusters met or overlapped. By reducing the number of nonconforming pixels and the amount of overlap, confusion between cover types was reduced at the expense of leaving more of the image unclassified. This procedure was called training area purification.

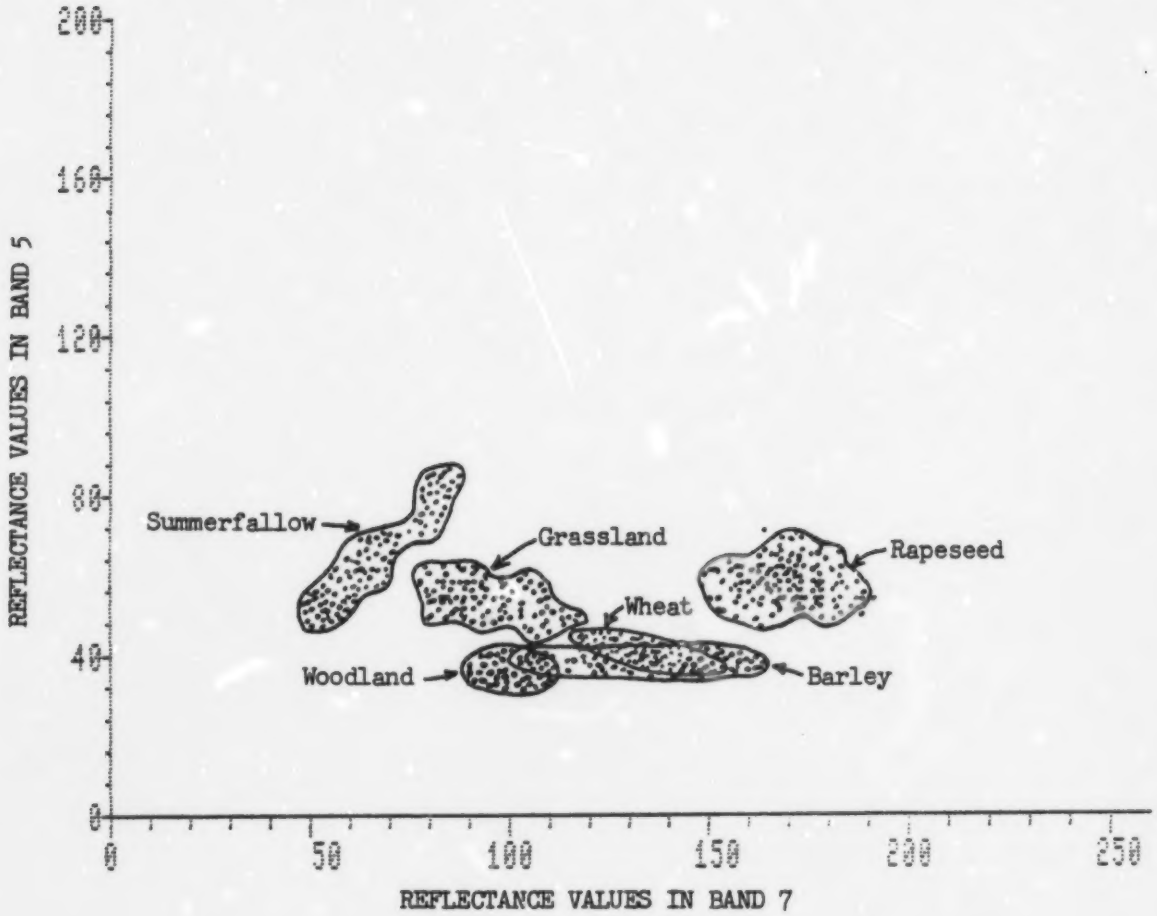


Figure 2. Scattergram showing the pattern and location of spectral clusters associated with six training sets for image TC1.

Training area purification was a process whereby a maximum likelihood thresholding function was applied to all pixels in a training set. Only those pixels which fell within an operator defined percentage of the mean were retained to be used later for signature generation. After each purification run, the scattergram with the reduced data set was viewed and if necessary the purification process was conducted again until the number of nonconforming pixels and overlap between spectral clusters was reduced. It is important to note that the scattergram represents only two feature files and that the actual signatures for each cover type were generated from a total of four feature files. Therefore, two clusters that appeared to overlap in two dimensional vector space may in fact be better isolated in four dimensional vector space.

When all training areas had been purified to a satisfactory degree, a spectral signature for each training set was generated under program TS. These signatures were then used as input into the maximum likelihood classifier running under program CB. The classification program was run in batch mode and depending on the size of the subarea required 1 to 4 hours of processing time.

Results of the initial classification were displayed under program IT and segment training areas were reloaded for easy reference. A visual comparison was made between ground reference data and classified results for each segment. In making this comparison attention was focused on several items. The number and distribution of unclassified pixels were noted. If large blocks of pixels remained (unclassified) in target crop areas after the first attempt, new training areas were created to include the unclassified pixels. Extreme caution is recommended in selecting

pixels for retraining since confusion with other cover types could result in further misclassification. In order to reduce the chance of this happening each time new pixels were added to the training set, the scattergram was examined and an assessment was made whether to retain or discard the new training site.

Misclassified pixels were also important to note during the preliminary assessment. Modifications to training sets were either in the form of an expanded or reduced signature file. For example, if the signature for canola resulted in classification of all the canola fields as well as say for example, some of the cereals, then the signature file for canola was reduced in the area of cluster which overlapped with the cereal's cluster. When all retraining and signature modification was complete, another supervised classification was run.

To determine whether retraining and reclassification had increased the classification accuracy, the area of each cover type was compared on a segment basis with the new pixel count. By comparing these areas to the known acreages of each cover type, the accuracy of the classification could be quickly estimated.

2.4.3 Post Classification Analysis

Several tasks comprised the post classification phase of image analysis. These tasks resulted in a filtered image, area measurements, accuracy assessments and generation of maps. Some of the tasks involved steps that were either operator or CPU intensive. As a result, several steps in these tasks could be performed simultaneously in order to maximize utilization of the system.

A minimum area filter or custerer which ran under program "CU" enabled analysts to clean up the "salt and pepper" effect of thematic maps. The filter permitted replacement of small areas with larger surrounding areas according to operator assigned minimum areas and relative importance of each class. The filtering process consisted of two stages both of which required minimal operator time to initiate once the filter parameters had been established. As a result, the filter process could be run in the background while at the same time performing other tasks.

The first stage of filtering was the map generation process where each pixel was assigned a value equal to the size of the contiguous area to which it belonged. All those pixels that could not be connected to a defined contiguous block were flagged as replaceable. At the completion of each map generation process the system responded with a suggested eat-in depth. If this depth was less than 7 the analyst could not choose to exceed the value, but instead had to choose an eat-in value that was less than or equal to the default value. If the default value was insufficient to completely replace all replaceable pixels then a second or subsequent passes were required.

The second stage of filtering was the replacement process where pixels flagged as replaceable were assigned to a larger neighbouring class. The analyst had the option of selecting the minimum area size to be replaced and relative importance of each theme. The replacement process could "eat-in" up to a maximum depth of 7 pixels. This meant that in order to replace a pixel that had been flagged as replaceable the filter searched for this area from all directions through a continguous block up to a maximum depth of 7 pixels. Therefore, a pothole in the middle of a large field would only be replaced if it was less than 7 pixels from the field's border, providing the rest of the field had been classified as the same theme (i.e. remained contiguous).

Parameters specified to create a filtered image for image SWQ were as follows:

STEP ONE: MAP FILE GENERATION

1. Diagonally connected pixels were considered to be in the same area.
2. Smallest undefined area was 9 pixels.
3. Default eat-in depth was 3 pixels.

STEP TWO: REPLACE MINIMUM AREAS

1. Maximum eat-in depth was 2 pixels.
2. Each theme was given the same weight when a replacement theme was selected.
3. Different minimum areas assigned to each theme were as follows:

Theme Number	Description	Minimum Area (Pixels)
1	Rapeseed	2
2	Summerfallow	2
3	Woodland	1
4	Grassland	5
5	Barley	10
6	Sunflower	8
7	Wheat	3
8	Water	1

4. Unclassified pixels were treated as a theme and assigned a minimum area of 9 pixels.
5. Two passes were required to generate the desired results.

The parameters specified to create the filtered image for image TC1 were as follows:

STEP ONE: MAP FILE GENERATION

1. Diagonally connected pixels were considered to be in the same area.
2. Smallest undefined area was 9 pixels.
3. The default eat-in depth was 7 pixels.

STEP TWO: REPLACE MINIMUM AREAS

1. Maximum eat-in depth was 6 pixels.
2. Each theme was assigned the same weight when a replacement theme was selected.
3. Different minimum areas assigned to each theme were as follows:

Theme Number	Description	Minimum Area (Pixels)
1	Rapeseed	3
2	Summerfallow	2
3	Sunflower	6
4	Flaxseed	4
5	Barley	4
6	Wheat	3
7	Water	1
8	Grassland	7
9	Woodland	2

4. Unclassified pixels were treated as a theme and assigned a minimum area of 5 pixels.
5. Only one pass of the filter was required to achieve the desired effect.

The parameters specified to create the filtered image for RIM were as follows:

STEP ONE: MAP FILE GENERATION

1. Diagonally connected pixels were considered to be in the same area.
2. Smallest undefined area was 10 pixels.
3. Default eat-in depth was 4 pixels.

STEP TWO: REPLACE MINIMUM AREAS

1. Maximum eat-in depth was 3 pixels.
2. Each theme was given the same weight when a replacement theme was selected.
3. Different minimum areas assigned to each theme were as follows:

Theme Number	Cover Type Description	Area (Pixels) Minimum
1	Rapeseed/Canola	3
2	Mixed Grassland/Scrub	3
3	Water	2
4	Grassland/Pasture	4
5	Woodland/Conifer	2
6	Corn	4
7	Cereals	8
8	Exposed Soils	3
9	Woodland/Mixed	4
10	Potatoes	3

4. Unclassified pixels were treated as a theme and assigned a minimum of area of 9.

5. Only one pass of the filter was required.

The minimum size areas selected for the filter were based on a visual comparison between classification results and ground reference data. Several combinations of minimum size areas were tried; however, these parameters were found to yield the most satisfactory resutls. No filtered image was generated for image CDM.

While the filter parameters were being processed, pixel tabulations were performed using the program "CT". The process was operator intensive because it was necessary to respond to several system prompts and specify the four character area name for each segment. Modifications to this program that would eliminate the need to repeatedly respond to system prompts and that would enable input of a list of training areas would reduce operator interaction time considerably.

Pixel tabulations were made on both a segment and crop district basis for both filtered and unfiltered results. Output from this process was in the form of numbers of pixels and hectares per theme for the land area contained within the specified training area.

Accuracy assessment of classified results involved the identification of problem segments following satellite image processing. Any significant variability between satellite and reported NFS survey data was referred to as an outlier. Data plots were used to readily identify these outliers by plotting NFS survey results (in acres) at the segment level against raw satellite classification results (in pixels)at the segment level, also.

A separate data plot was run for each CCT and crop of interest. Problem segments were then cross-checked with survey results as recorded on the 'enlarged' aerial photographs as well as to the survey questionnaire data. If the discrepancy between the NFS and the satellite data remained unresolved following this step, further investigation required on-site verification of the problem area and/or a telephone follow-up to the NFS respondent. Following identification of the error source, appropriate corrective measures were carried out.

Although the desired output from this project was statistical data, a sample of the classification results in the form of a thematic map is found in Appendix II. Thematic data were resampled under program "RS" using the nearest neighbour resampling algorithm. The resampled image was then reformatted using program "OA" so that classification results could be reproduced as a color thematic map on an Applicon plotter. Both the filtered and unfiltered results for image TC1 which corresponds to the eastern 1/3 of crop district 1 were produced.

2.5 Crop Area Estimation

2.5.1 Post Stratification

For remote sensing purposes the area frame was post stratified according to the variable of interest. The NFS area frame was originally stratified according to total cattle and total cropland after removing the list farms. It was felt that this stratification would not be optimal for remote sensing purposes. Therefore, post stratification was based on area of rapeseed as determined from the 1981 Census of Agriculture and only three strata were defined.

The rationale for selection of three strata was as follows. The actual NFS design usually resulted in formation of many strata (sometimes more than 10) for each SPA. Many of these strata have a small sample size and often each strata had a sample size that was as low as two EA's. Loss of Landsat data for entire EA's and/or portions of others during the optimum period for crop discrimination due to cloud cover was not an unusual event. Therefore three strata were chosen in order to avoid a small sample size. In addition, use of the regression estimation technique can be viewed as a means of stratifying because it makes use of existing auxillary information.

It is important to note that before proceeding to the post stratification, new EA summaries were arrived at by summing all the farms included in the concerned EA's without exclusion of any farms.

2.5.2 Regression Estimation

A regression estimator was used to obtain the total acreage of the target cover types (rapeseed, cereals and summerfallow). Estimates along with their variance were calculated using the NFS post stratified design. The EA's were treated as a with-replacement-sample. Segments and EA's with missing satellite data were not included in the sample, nor calculated as part of the population in order to produce first estimates for the unclouded area of SPA1. Areas within segments where major discrepancies occurred (i.e. outliers, non-contacts and refusals) that could not be resolved by follow-up methods were either removed or imputed.

A second set of estimates for the entire population (SPA1) was produced using the first estimates and multiplying these estimates by the ratio of the total acreage of the crops of interest for the entire SPA at 1981 Census over the total acreage of the crops of interest of the unclouded portion of SPA1 at 1981 Census.

The formulae of the regression estimator which utilizes both ground data and classified Landsat pixels was as follows. Let the label X represent the reported satellite crop of interest (in pixels) and Y represent the ground data of the crop of interest (in acres). M_{hi} was the total number of segments for the selected ith enumerative area in stratum h. The m_{hi} was the total number of selected segments for the selected ith enumerative area in stratum h. M_h was the total number of enumerative areas in stratum h. The total number of selected enumerative areas in stratum h was represented by n_h. The regression estimate of Y was then calculated as:

$$Y_{Reg.} = \hat{Y} + B(X-\hat{X}),$$

where $$\hat{Y}=\sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} y_{hij}$$

was the design estimate of Y; $\hat{X}$ was the design estimate of X, the total uncorrected satellite target crop area in the test region. $\hat{X}$ was obtained by substituting x_{hij} for y_{hij} in the formula for $\hat{Y}$; y_{hij} and x_{hij} were the observed values for the jth selected segment in the ith selected first-stage unit (EA) of stratum h; $B=\text{cov}(\hat{Y},\hat{X})/\text{var}(\hat{X})$ was the linear regression coefficient.

The variance estimate and $\hat{Y}_{Reg.}$ was given by:

$$v(\hat{Y}_{Reg.}) = v(\hat{Y}) - 2\hat{B}\,\text{cov}(\hat{Y},\hat{X}) + \hat{B}^2 v(\hat{X}),$$

where $$v(\hat{Y}) = \sum_{h=1}^{L} \frac{N^2_h}{n_h(n_h-1)} \sum_{i=1}^{n_h} \left[\frac{M_{hi}}{m_{hi}} \sum_{n=1}^{m_{hi}} y_{hij} - n_h^{-1} \sum_{i=1}^{n_h} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{m_{hi}} y_{hi} \right.$$

and $v(\hat{X})$ was calculated by subsituting x_{hij} for y_{hij} in $v(\hat{Y})$. It can be seen that $\hat{B} = \text{cov}(\hat{Y},\hat{X})/v(\hat{X})$ is the value of B which minimizes $v(\hat{Y}_{Reg.})$.

3. RESULTS

Although the entire study region was covered by satellite imagery, area estimates for each crop district were not possible. Because of differences in image dates and radiometric characteristics of the satellite data, a separate regression was required for each Landsat scene to produce a total estimate for all target crops in SPA1 (Table 2). Only results of unfiltered data were shown because filtered results were not generated for image CDM or RIM and subsequently not for SPA1. Because of difficulties with ground reference data and too few segments, a separate regression analysis for image RIM was not calculated for either filtered or unfiltered data.

Crop district 3 was covered by one image (CDM); however, there was a clouded area that covered 25.2% of the total rapeseed acreage of this district as reported in the 1981 Census of Agriculture. The rapeseed estimate for the unclouded area was calculated as 92,549 acres (c.v.= 7.7%). When expanded up to the crop district level, the remote sensing estimate became 124,720 acres (c.v.= 7.7%).

The estimate for SPA1 was obtained by adding the crop district 3 estimate with the estimates for iamges TC1 and SWQ. To take into account the production of the area covered by image RIM which represented 8.55 percent of the total rapeseed acreage of SPA1 at the 1981 Census, the previous estimate was multiplied by a factor of 10000/9145. This factor represented the ratio of total rapeseed acreage of SPA1 over the rapeseed acreage of the area covered by crop district 3 and images TC1 and SWQ at the 1981 Census. The same factoring technique was applied for the cereal grains and summerfallow.

TABLE 2

REMOTE SENSING REGRESSION ESTIMATES FOR SPA 1 AND EACH OF THE SATELLITE IMAGES

CROP	IMAGE	ESTIMATED ACRES (C.V)
Rapeseed	[1]CDM	124,720 (7.7%)
	SWQ	115,580 (9.2%)
	TC1	27,182 (13.4%)
	[2]RIM	25,026 N/C
	SPA 1 Total	292,508 (5.5%)
Summerfallow	[1]CDM	362,232 (13.8%)
	SWQ	210,310 (15.1%)
	TC 1	36,930 (20.3%)
	[2]RIM	38,738 N/C
	SPA 1 Total	648,210 (9.8%)
Cereal Grains	[1]CDM	956,337 (7.7%)
	SWQ	1,336,013 (5.7%)
	TC 1	289,516 (10.3%)
	[2]RIM	303,679 N/C
	SPA 1 Total	2,885,545 (4.3%)

FOOTNOTES

1. Estimate for image CDM has been adjusted to account for cloud cover in Crop District 3.

2. Insufficient segment data to generate a remote sensing regression estimate for image RIM. An approximation for image RIM was given by subtracting the total of images CDM, SWQ and TC 1 from the SPA1 total. An estimate for the coefficient of variation was not calculated (n/c).

The coefficient of variation for each of the remote sensing regression estimates was always greater than the coefficient of variance for the SPA1 estimate. This difference can be expected because of the fewer samples and relatively smaller area associated with each satellite image as compared to the total for the SPA.

Target crop area estimates using remote sensing numbers were calculated at the crop district level (Table 3). It is important to note that estimates for crop districts 1 and 2 were only approximations calculated with the aid of the total pixel count for each of the three images (Appendix III). These estimates were not subjected to the same rigorous statistical procedures as for crop district 3 and SPA1 but do represent the only approximation for remote sensing at the crop district level.

The difficulty in calculating the value for crop district 1 and 2 was the result of having calculated only one remote sensing regression estimate for image SWQ which covered both districts. It was not possible to run two separate regressions for the two crop districts within this image because there were too few segments. Therefore, allotment of the original remote sensing regression estimate (115,580 acres) to each crop district was based on the ratio of target crop pixels in each crop district. For example, there were a total of 89,208 rapeseed pixels for image SWQ of which 47,477 were counted in crop district 1 and the remaining 41,731 were in crop district 2.

The ratio of crop district 1 pixels over the total pixels for image SWQ multiplied by the remote sensing regression estimate gave an approximate rapeseed acreage for that portion of image SWQ which corresponded to crop district 1. The total acreage of rapeseed in crop district 1 was then calculated

TABLE 3

1983 MANITOBA CROP MONITORING REMOTE SENSING ESTIMATES FOR CROP DISTRICTS 1, 2 and 3

	CROP DISTRICT ESTIMATES Shown in Acres (C.V.)			TOTAL ACRES
	1	2	3	
Rapeseed	88,694	79,094	124,720 (7.7%)	292,508 (5.5%)
Summerfallow	136,717	149,261	326,232 (13.8%)	648,210 (9.8%)
Cereal Grains	869,396	1,059,812	956,337 (7.7%)	2,885,545 (4.3%)

by adding the estimated value for image SWQ and the calculated remote sensing value from image TC1. The estimate for crop district 2 was calculated in the same manner except the estimated value for rapeseed from image RIM was added to the value of the SWQ component. As mentioned previously, crop district 3 estimates were generated from image CDM. Crop district estimates for cereal grains and summerfallow were also calculated using this method.

4. DISCUSSION

4.1 Area Estimation

In comparing the remote sensing regression estimates at the SPA level with results of other area estimation techniques, confidence intervals were within a very acceptable range for both rapeseed and summerfallow (Table 4). As might be expected, the remote sensing estimate for cereal grains was significantly higher than the estimates produced by the NFS and Crop Reporting techniques.

The remote sensing estimate for rapeseed had a lower coefficient of variance (c.v.= 5.5%) than the NFS estimate (c.v.= 12.9%) but still fell well within the range of the NFS estimate. Although the upper limit of the range for the remote sensing estimate was less than the provisional Crop Reporting estimate, the difference was not considered to be significant.

Slightly higher remote sensing estimates for rapeseed likely could have been generated if the rapeseed training areas had not been purified. Training area purification and hence reduction in the range of the spectral signature for rapeseed resulted in fewer pixels being classified as rapeseed. Although not quantified, this observation was supported by a visual comparison between segment data and classified results for image SWQ (Appendix IV). On this image, segments 6, 9, 12, 24 and 53 had numerous unclassified pixels associated with canola fields. Image classifications performed with non-purified rapeseed training areas produced such poor classification results that further statistical analysis was not recommended. The results finally arrived at through training area purification were a compromise between over purified and non-purified training areas.

TABLE 4

1983 MANITOBA CROP MONITORING
STUDY SURVEY ESTIMATES (ACRES)

CROP	SURVEY	*CROP DISTRICT 1 & 2	CROP DISTRICT 3	SPA1
Rapeseed	[1]Remote Sensing (regression)	167,788 n/c	124,720 c.v. (7.7%)	292,508 c.v.(5.5%)
	[2]N.F.S. (MTF)	n/c	n/c	294,865 c.v.(12.9%)
	[3]Crop Reporting	202.800 c.v.(n/a)	124,300 c.v(n/a)	327,100 c.v. (n/a)
	Census of Agriculture	-	-	201,188
Cereal Grains	[4]Remote Sensing (regression)	1,929,208	956,337 c.v.(7.7%)	r2,885,545 c.v.(4.3%)
	N.F.S. (MTF)	n/c	n/c	2,555,717 c.v.(4.4%)
	[5]Crop Reporting	1,713,700 c.v. (n/a)	846,100 c.v. (n/a)	2,559,800 c.v.(n/a)
	[6]Census of Agriculture	-	-	2,527,015
Summerfallow	Remote Sensing (regression)	285,978	362,232 (13.8%)	r648,210 c.v.(9.8%)
	N.F.S. (MTF)	n/c	n/c	640,827 c.v.(6.3%)
	Crop Reporting	302,000 c.v. (n/a)	260,000 c.v. (n/a)	562,000 c.v (n/a)
	Census of Agriculture	-	-	614,245

FOOTNOTES

(*) A separate estimate for Crop Districts 1 and 2 was not generated for the remote sensing survey but was calculated by subtracting the estimate for Crop District 3 from SPA1.

(n/c) Not Calculated.

(n/a) Not Applicable.

(r) Revised February 27, 1984.

(1) Estimates are for unfiltered data only.

(2) Only MTF estimate of NFS generated at SPA1 level.

(3) Crop Reporting Panel - NOVEMBER ESTIMATE.

(4) Remote Sensing Regression-Cereal Grains include total wheat, oats, barley, rye, mixed grains and oats for fodder.

(5) Crop Report Panel - Total Grains include total wheat, oats, barley, total rye and mixed grains.

(6) 1981 Census of Agriculture - Total grains include wheat, oats, barley, total rye, mixed grains and oats for fodder.

The general tendency of the remote sensing estimate for rapeseed to be lower than the other estimated values could also perhaps be related to imagery dates and crop conditions. Analysis of images acquired on July 14 (CDM and SWQ) may not have yielded results for all the rapeseed as only 80 to 90 percent was in flower. Rapeseed was distinguished from other crops when it was in flower. Furthermore, analysis of the image acquired on August 8 (TC1) could have missed some rapeseed because harvest had just begun (less than 5 percent).

As with rapeseed, the confidence interval for summerfallow was well within acceptable limits of the NFS estimate. The lower end of the range for the remote sensing estimate was greater than the Crop Reporting estimate but the difference was not considered significant.

One explanation for a generally high remote sensing estimate for summerfallow was the fact that in some instances road allowances and other areas of sparse vegetation cover were classified as summerfallow. The spectral signature of the later two categories were very similar to summerfallow and even extensive training area purification could not alleviate this problem of overlap between fallow and roads. It is important to note that not all road allowances were classified as summerfallow. In fact, road allowances often showed up as other crop types or as natural grasslands depending on the positioning of the pixel and the composite spectral characteristics of the ground features making up that pixel. In addition, many areas of natural grasslands that were overgrazed or which were sparsely vegetated such as those areas associated with the Stocton Loamy Sands in the Carberry Hills were classified as summerfallow. Therefore, not only did the final remote sensing regression estimate contain summerfallow fields but also portions of road

allowances and sparsely vegetated areas. This would explain the slightly higher results for the remote sensing values, especially given that these cover types were not always included in segments since they often were the features at the segment borders and may not have been accounted for in the regression.

The remote sensing estimate for cereal grains was significantly higher than the NFS and Crop Reporting Estimates. Several factors contributed to this overestimate; however, the two most significant factors were a broad spectral signature for the cover type and confusion with grasslands, pastures and sparsely vegetated shrublands due to signature overlap. As a result of sample design, these other crops were not included in the segments in sufficient quantity to have been removed during the regression stage.

The cereal grain category was comprised of several grains including wheat, oats, barley, rye, mixed grains and oats for fodder and this in itself presented problems. With this mix of crops in the cereal cover type it was not unusual to find a large and highly variable range in the spectral signature. Variable stages of crop development and the influence of numerous site-specific conditions (i.e. soils, slope, crop density) all worked together to influence the spectral signature. Extensive training area purification could not resolve the problem of a broad spectral signature for cereal grains.

Another factor affecting the accuracy of the cereal grain estimate was overlap in the spectral signature between cereal grains and other forms of grasses. Natural grasslands, pastures and many forage crops displayed similar spectral characteristics to the cereal grains. Other cover types which displayed similar spectral characteristics to the cereals were flax, corn and sunflower. One might suspect that given the significant differences in physical appearance and phenology of these crops that they could be separated from the cereals. In order to attempt such a separation satellite imagery would have to be acquired at a time when these differences were at a maximum.

4.2 Sample Design

The NFS sample design did not exactly meet remote sensing needs for two major reasons. First, those farms used to build the list frame were removed from the area frame with the result being a non optimal stratification for remote sensing. The key variables used to build the list frame were total cattle, pigs, cropland and sales. Secondly, stratification of EA's using total cattle and total cropland as key variables in the hierarchical cluster analysis was not optimumal for remote sensing purposes. The result of differences between remote sensing needs and the survey design was that regression estimates were not possible for crop districts 1 and 2 or for image RIM because segment locations were not optimum.

Although the present method of segment location was not optimal for remote sensing regression needs, the type of ground reference data gathered as part of the NFS was extremely well-suited to interpretation and verification of Landsat data. The satellite data could not have been interpreted without information on crop type, its exact location and field boundaries. In addition to this information, an optimum design for a remote sensing project should consider Landsat frame boundaries and placement of segments in relation to soil types.

Because radiometric calibrations to the Landsat data were performed on a scene basis, the performance of the maximum likelihood classifier varied from scene to scene. Therefore a separate regression analysis was conducted for each scene. Furthermore, classifier performance within a scene was affected by major changes in soil conditions necessitating further stratification considerations.

During this project, a spectral signature for one crop type could not always be applied to the same crop type located on the same Landsat scene but in a different soil zone. In other words the influence of soils on crop

development, as well as background reflectance characteristics of the various soils was significant to the extent that separate training areas and signature development were required for each major soil zone. These findings were consistent with Schubert and Chagarlamudi (1978) who found that the most significant factors affecting wheat productivity were soil-climatic factors. Table 5 shows their proposed hierarchy of soil factors which affected wheat yields. An optimum remote sensing survey design must consider the differences in classifier performance between Landsat frames and should base stratification within each frame on a hierarchy of soil factors.

4.3 Post Classification Filtering

Application of a post classification filter was an extremely time consuming (operator and CPU) operation that did not necessarily result in production of better statistics. The filter did however result in production of an obviously much cleaner-looking map product (Appendix II).

Part of the problem with the filter was that selection of its parameters was somewhat subjective. Furthermore, Ryerson and Thibault (Pers. Comm.) suggested that the regression analysis and stratification procedures compensated for minor discrepancies in the classification.

The subjectivity of the filtering process resulted from the assignment of minimum areas (i.e. weighting factors) to each class. The trial and error process of assigning weighting factors was based on the contiguity of classified results in relation to ground reference data, class type and intrusion pattern of incorrect pixel classifications into known target crops and land use patterns characteristic of the target crop.

TABLE 5

HIERARCHY OF SOIL FACTORS WHICH AFFECT WHEAT YIELDS (Schubert and Chagarlamudi 1978).

1. SOIL - CLIMATIC FACTORS

   Present Climate
   (climatic zone, subregion)

   Previous Climate
   (soil morphology)

2. TEXTURE

3. SOIL PARENT MATERIALS

   Mode of Deposition
   (aeolian, fluvial, glacial)

   Composition
   (till, boulder, cretaceous shale)

   Residuals
   (stoniness, alkalinity, salinity)

4. TOPOGRAPHY

   Slope

   Surface Expression
   (landform, drainage)

5. LOCAL EFFECTS AND EFFECTS OF USE

   Irrigation

   Wind Erosion

Contiguity referred to the "salt and pepper" appearance of the classification. In otherwords, were the classification results in large continguous blocks of the correct theme for each target crop or were the results spotty? If the results were spotty, which classes were most spotty and what was the average size of the contiguous spots (i.e. 1 to 3 pixels, 4 to 6 pixels, etc.)? Next intrusion patterns were examined to determine which themes were intruded into target crops and what was the average contiguous size of these intrusions (i.e. 1 to 3 pixels, 4 to 6 pixels, etc.). Land use patterns were also considered in assigning weight factors. For example, grain fields in southwestern Manitoba were usually quite large (i.e. $> \frac{1}{4}$ section) therefore by specifying a large minimum area for replacement, say 7 to 10 pixels, there would be little chance to eliminating an entire wheat field. On the otherhand, minimum areas for rapeseed were kept quite small because the likelihood of smaller (i.e. $< \frac{1}{4}$ section) rapeseed operations was quite high. Subsequently, there was a high risk of filtering out a small rapeseed operation.

A fixed parameter, box car filter (3 x 3 pixels) would be a useful feature for the filtering process.

4.4 Boundary Placement

Difficulties with boundary placement for segments presented additional sources of error. Records from previous projects indicated that boundary placement with uncorrected data produced segments of highly variable size. Results for this project showed somewhat less variable segment sizes; however, the significance of this variation was noted (Appendix IV). For example, on image SWQ pixel counts for rapeseed could have been altered in segments 12, 34, 36, 37, 38, 54 and 55 if the segment border had been shifted by as little as one pixel.

The problem related in part to the location of the start or corner pixel, orientation of the segment and the skew in the satellite data. For example, if the pixel used as the inside corner fell exactly on the northwest corner of the northwest field, the result would be different than if a pixel on the inside corner fell partly on a road and partly on a field. The effect of this difference in pixel placement was to either include or exclude highly variable border pixels in the segment.

The whole issue of boundary placement was further complicated by variations in nominal pixel size and skew in the satellite data. As a result of skewed data, north-south and east-west straight lines were approximated by a "stepped" line. The effect of data skewness was to limit accuracy in placement of segment boundaries and to arbitrarily include or exclude pixels on the segment borders.

A simplistic assessment of the boundary problem could be made by assuming no skew problem, constant pixel size and a north-south oriented segment with the northeast corner pixel positioned exactly in the corner. Given a nominal 59 m resolution across track, the northern or southern segment boundary (1609 metres) would contain 27.3 pixels. The mixed pixel at the northeast corner would be missed only $\frac{1}{4}$ of it. The segment width would be approximately 27 pixels. If the mixed pixel at the start included road, roadside and crop then there could be a start pixel composed of 0.3 road, 0.4 grass and 0.3 crop (or 23.6 metres of the 79 metre collection pixels). Remaining from the 1609 metres would be 1585.4 metres or 26.9 pixels. The sensor delay or spill-over effect from the crop to the road could easily translate to an additional pixel. With a north south segment length of

4827 metres (61.1 pixels with 79 m resolution) the segment could gain 61 to 122 pixels. With the stepped line effect this could translate into considerably more. Furthermore, a similar gain (or loss) of 27 to 54 pixels in the east-west direction could be expected. With the total number of border pixels at 176, it would not be unreasonable to expect variability in segment size from 88 to 176 pixels from the norm of 1647 pixels (N-S oriented). East to west oriented segments would contain 20.3 X 81.8 pixels - say 20 X 82 - or 1640 pixels. The border would, however, contain 204 pixels.

Errors of boundary placement could be reduced significantly by using geometrically corrected data. Although the location of the corner pixel could still fall partly on a road, the effect of the stepped line would be eliminated. Geometric correction was a time consuming process which would have to be weighed carefully against the other advantages of using geometrically corrected data. Perhaps the greatest advantage would be the ability to interface a digitizer with the satellite data. This feature alone could save numerous hours of operator interaction time required to locate segment, soil, municipality, crop district and between scene boundaries.

4.5 Feasibility of a Provincial Crop Monitoring Program

Several important questions must be answered in order to determine whether a provincial crop monitoring program using satellite data is feasible. These questions concern the need for provincial crop information, technical considerations, facility and staffing requirements, coordination and implementation of a program.

While a detailed discussion of the need for provincial crop information was beyond the scope of this paper it will be sufficient to say that this question has likely already been aswered. In the Manitoba Agriculture

Year-Book (1981) with the value of crop production reported at just over $1,451 million and with the total area of crop production estimated at 11 million acres (4 million hectares) the question really becomes how can the provincial crop reporting system be improved so as to increase returns from sales on domestic and international markets. Clough (1974) reported that by reducing errors in the Canadian wheat production forecasts by just 1 to 2 percentage points the net gain of export revenue would be about 13 to 22 million dollars per annum and the net gain of producers' surplus would be about 6 to 11 million. These estimated benefits alone should be large enough to justify a substantial program to improve domestic production forecasting methods.

Improvements in crop forecasting by using satellite data will be achieved when technical questions concerning crop separability, yield determination, data availability and analysis methods have been answered. Ryerson et al (1981) stated that significant advancements have already been made in these areas and the encouraging results achieved in this project provided further support to this statement.

It is not unreasonable to expect that to conduct a provincial crop monitoring program over an area approximately three times the size of the 1983 regional project, that it would require about three times the level of effort. Assuming a 24-hour operating day during the peak analysis period (August and September) there were a possible total of 1488 operating hours for one ARIES-II system and double this amount (2976 hours) if two image analysis systems were available. Considering that the total amount of time utilized by both systems (Winnipeg and Ottawa) for the 1983 project was

approximately 550 hours including developmental efforts and some training there is ample room for increasing total demand on the computer facilities by a factor of 3 and still maintain a very comfortable safety margin. Furthermore, improvements to the system hardware and software which have already been realized in Winnipeg and operator experience gained from this project would likely mean even faster turnaround times for data processing. Clearly, the demand for image processing facilities for a provincial program could easily be accommodated.

The level of staffing required to conduct a provincial crop monitoring program was estimated with less certainty because of the number of agencies involved. Two elements working in favour of meeting staffing requirements for an expanded program are the skill and expertise of each agency , the existing infrastructure and the knowledge that these agencies can coordinate their activities as demonstrated in this project. With adequate lead time, careful planning and continued interagency cooperation a provincial crop monitoring program would be feasible.

5. CONCLUSIONS

Results of the 1983 Crop Monitoring project have shown that remote sensing techniques can be used to produce reliable area estimates for rapeseed, summerfallow and cereal grains at the sub-provincial level in Manitoba. The confidence interval of the rapeseed and summerfallow estimates were within acceptable limits of the National Farm Survey and Crop Reporting estimates. Remote sensing estimates for cereal grains were significantly larger than the Crop Reporting estimate but fell within the limits of the National Farm Survey estimate.

Differences in values between remote sensing and other estimation techniques were attributed to several factors. With rapeseed there was a general tendency to underestimate this class. Training area purification and hence reduction in the spectral signature for rapeseed accounted for some of this loss. Further underestimation resulted from differences in image dates and crop conditions. Images acquired on July 14 found rapeseed only 80 to 90 percent in flower whereas images acquired August 8 found just under 5 percent of the rapeseed already harvested.

Overestimation of summerfallow and cereal grains was attributed to confusion with other cover types not adequately represented in the sample. Roads and sparsely vegetated areas were included with the summerfallow, whereas partial inclusion of confusion crops (i.e. natural grassland, improved pasture, flax and corn) into the cereal grains resulted in overestimation for this category.

Although the ground reference data gathered as part of the NFS survey was well suited to interpretation of Landsat data, the survey design was not optimal for remote sensing needs. The present NFS survey had a

multipurpose, multiframe design which used total cattle and total cropland as the key variables for stratification. A survey design that considered the juxtaposition of Landsat frame boundaries in relation to a hierarchy of soil factors would result in better placement of segments and a survey design more in line with remote sensing requirements.

Post classification filtering was a time consuming operation that resulted in production of good looking map products; however, it did not result in generation of area estimates that were any better than unfiltered classification. One of the major problems with the filter was the subjectivity involved in selecting filter parameters. Furthermore, regression analysis on unfiltered data accounted for many of the minor discrepancies in classification results and analysis of unfiltered data yielded satisfactory statistical values.

Problems with boundary placement produced segments which could vary in size from 88 to 176 pixels from the norm of 1647 pixels for a north-south oriented segment. Geometric correction of satellite data would eliminate the "stepped-line" effect on segment boundaries which would reduce variability in segment size.

A provincial monitoring program for rapeseed, summerfallow and cereal grains was considered to be technically feasible and could be accomplished with existing facilities and staff. The success of such a program would depend on careful planning and close cooperation between federal and provincial agencies.

6. RECOMMENDATIONS

This project has demonstrated that the infrastructure needed to handle both technical and organizational aspects of domestic crop reporting using remote sensing techniques, now exists at the level required for regional crop reporting. Before expanding this program to a provincial level, careful consideration should be given to the following recommendations:

1. The survey design and segment location process should be designed to meet remote sensing requirements. Stratification within Landsat scenes should be related to soil factors. In order to accomplish this, additional segments would have to be added to the existing NFS sample.

2. A standardized procedure should be developed for placement of segment boundaries.

3. Further analysis of post classification filtering is required to determine how this procedure can be used to improve area estimates.

4. The feasibility of using geometrically corrected CCT's should be examined.

5. An interagency agreement should be established and maintained to ensure a corrdinated approach in the design, implementation, and release of information from a provincial crop monitoring program.

LITERATURE CITED

Brown, R.J., F.J. Ahern, R.A. Ryerson, K. Thompson, D. Goodenough, J. McCormick and P. Teillet. 1980. Rapeseed: Guidelines for Operational Monitoring. 6th Canadian Symposium on Remote Sensing, Halifax, N.S.: 585-596.

Chatterjee, S. 1967. A Note on Optimum Allocation. Skand. Akt., Vol. 50: 40-44.

Clough, D.J. 1974. Remote Sensing and Economic Values of Errors in Wheat Forecasting. A technical study of the Canada Centre for Remote Sensing, Dept. of EMR presented at the Third Meeting of the Agricultural Working Group of the Third Meeting of the Agricultural Working Group of the Canadian Advisory Committee on Remote Sensing. October 24, 1974. Charlottetown, P.E.I.: 24pp.

Goodenough, D. and S. Shlien. 1974. Automatic Classification Methodology. Canada Centre for Remote Sensing, Research Report 74-1: 8 pp.

Hanuschak, G.A., R. Sigman, M.E. Craig, M. Ozga, R.C. Luebbe, P.W. Cook, D.D. Kleweno and C.E. Miller. 1980 Crop-Area Estimates from Landsat; Transition from Research and Development to Timely Results. IEEE Transactions on Geoscience and Remote Sensing 18(2): 160-166.

Ingram, S. and G. Davidson. 1983. Methods Used in Designing the National Farm Survey. Proceedings of the American Statistical Assocation, Toronto, Canada: 6 pp.

Manitoba Agriculture. 1981. Manitoba Agriculture Year book. Economic Analysis Branch, Winnipeg, Manitoba: 115 pp.

Ryerson, R.A., J.L. Tambay, R. Plourde, R.J. Brown, L.A. Murphy and B. McLaughlin. 1981. A Timely and Accurate Potato Acreage Estimate from Landsat: Results of a Demonstration. 15th Int. Symp. on Rem. Sens. of Environ., Ann Arbor, Mi: 587-597.

Ryerson, R.A., R.S. Sigman and R.J. Brown. 1981. Satellite Remote Sensing for Domestic Crop Reporting in the United States and Canada: A Look to the Future. 7th Canadian Symposium on Remote Sensing, Winnipeg, Man.: 30-41.

Ryerson, R.A., J.L. Tambay, R. Plourde, and J. Harvie. 1983. The Use of Landsat, Ground Data and a Regression Estimator for Potato Area Estimation. CCRS Manual 83-1.

Schubert, J. and P. Chagarlamudi. 1978. Analysis of Long Term Average Wheat Yield Data for use in the Assessment of Uniform Productivity Areas and Model Land Units, in Assessment of Uniform Productivity Area and Model Land Units from Field Evaluation and Long Term Average Wheat Yield for Use in Analysis of Landsat Data. The Sibbald Group, Haskins and Sells Associates. Calgary, Alberta: 45 pp.

APPENDIX I

Interviewer's Instruction Manual

NATIONAL FARM SURVEY

REMOTE SENSING PROJECT 1983

A co-operative study between Statistics Canada, the Canada Centre for Remote Sensing, the Manitoba Remote Sensing Centre and the Manitoba Department of Agriculture.

REMOTE SENSING PROCEDURES

INTERVIEWER'S INSTRUCTIONS

HISTORY OF THE PROJECT

During the 1981 and 1982 Farm Enumerative Survey, Statistics Canada investigated the use of satellite images for estimating rapeseed Canola acreages. Interviewers working in several areas in Alberta and the Peace River Block of British Columbia were asked, for certain segments, to outline and identify the content and size of all fields on special aerial photographs. The identification of all crops in the selected segments was required because of the similarity between the crops on the satellite images. The information collected by the Interviewers was used to "tell" the computer where rapeseed fields were located. The computer could then locate these known rapeseed fields on the satellite images, and by searching for and identifying other fields which looked the same, it could accurately estimate rapeseed acreages.

This method of estimating crop acreages is being developed so that in the future we can reduce the amount of information which must be obtained from farm operators. Because of the past success with this project, it is being continued this year in Alberta and is being conducted for the first time in Manitoba.

REMOTE SENSING PROCEDURES

FIELD PROCEDURES

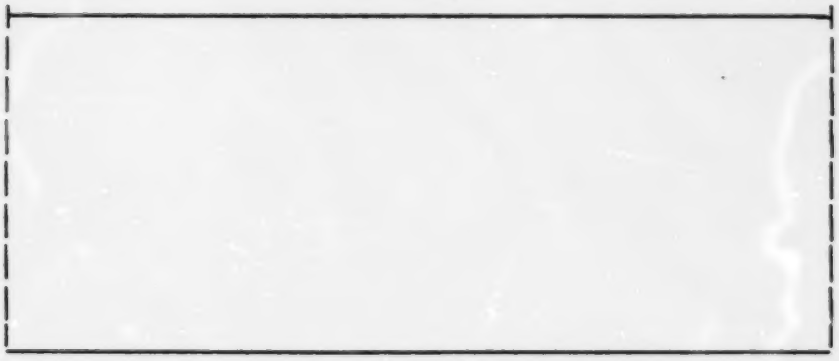
NOTE: THE FOLLOWING PROCEDURES ARE IN ADDITION TO THE PROCEDURES FOR THE NATIONAL FARM SURVEY, AND DO NOT REPLACE ANY PREVIOUS OR SUBSEQUENT INSTRUCTIONS.

You will find in your assignment package, special enlarged aerial photographs labelled "Remote Sensing". This identifies the segments which have been selected for the remote sensing test.

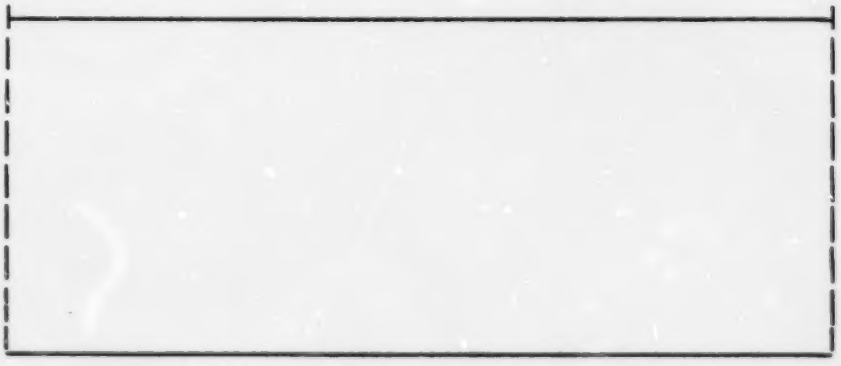
The following information is also included on the "Remote Sensing" photograph:

- the segment identification number;
- a direction label, indicating North;
- a 10 acre scale, and
- abbreviated codes as shown below (to be used as applicable):

LABEL 1:



LABEL 2:



REMOTE SENSING PROCEDURES

After you have completed Section D, "Land Use", on the NFS questionnaire, you will have a record of the type of crops on the operator's holding. Proceed at that time with the Remote Sensing procedures which are as follows:

- Give the respondent the Remote Sensing brochure to briefly explain the purpose of remote sensing.

- Ask the respondent to identify his/her holding within the segment on your remote sensing aerial photograph.

- Outline the boundaries of the holding with the special red marker provided to you.

- Ask the respondent to identify, in that holding, each parcel of land three acres or larger, according to its use.

- Outline the boundaries and identify the size of each field (in acres), using the codes shown on the aerial photograph and also included below:

GRAINS

B.....BARLEY
C.....CORN (for grain or fodder)
MG.....MIXED GRAINS (includes any mixture of wheat, oats, barley or rye)
O.....OATS
RY.....RYE
W.....WHEAT

OILSEEDS

F.....FLAXSEED
M.....MUSTARD SEED
RA.....RAPESEED
SN.....SUNFLOWER

REMOTE SENSING PROCEDURES

GRASSES	OTHERS
A.....ALFALFA	S.....SUMMERFALLOW
CL.....CLOVER	OC.....OTHER CROPS (e.g., peas, potatoes, buckwheat, vegetables, etc.)
G.....GRASS/PASTURE	OL.....OTHER LAND (e.g., barns, buildings, bushlands, creeks, headlands, sloughs, roads, etc.; applies to an area of a farm which is not used for agricultural purposes.
G.....GRASS/PASTURE	

There are other codes to identify holdings which were non-interviews, non-agricultural or other special situations as listed below.

NA ..... NON-AGRICULTURAL HOLDING;

NI ..... NON-INTERVIEW (to be used when no contact was made at the holding or for respondent refusals)

X ..... EXCLUSIONS (to be used for land which is part of holdings excluded from the NFS as outlined in Section 9.19 of the Interviewer's Manual, e.g., Canada Packers, Campbell Soup Co., Indian Reserves, etc.)

NOTE: **Where a type of grain is identified by 2 letters, be sure to enter both letters on the aerial photograph.**

- After outlining the fields, ask the operator the acreages of the fields in that parcel and record the response inside the field boundaries on the photo. If the operator reports, for example, that a certain parcel contains 160 acres of rapeseed and you can see some areas that appear to be non-productive, confirm that these areas are still non-productive areas.

If the total non-productive area is more than 3 acres, (for example, 15 acres), ask if those acres have been included in the estimate given for that parcel. If the non-productive acres have been included in the 160 acre estimate, you must adjust the total acreage on the aerial photograph accordingly:

RA	145
X	15

An estimate for non-productive land exceeding 3 acres is extremely important.

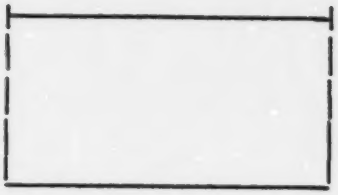
If the total non-productive area is less than 3 acres, report the total figure, i.e. RA 160.

- A square, representing 10 acres has been drawn on a corner of the enlarged aerial photo. Use the 10 acre square as a guide to judge the relative field sizes. For example, does the 40 acre field look twice as large as the 20 acre field? If the field sizes do not appear consistent, check:

  - that the farmer has assigned the acreage to the correct field;
  - that the farmer is referring to only the part of the field inside the segment;
  - that the field has not been enlarged since the photo was taken;
  - that the farmer is not referring to a group of fields.

REMOTE SENSING PROCEDURES

Field boundaries should be changed to account for any discrepancies.



- Outline all parcels as precisely as possible. For example, if the land use has changed since the aerial photo was taken, (e.g., a wooded area cleared and now used for crops) you would outline the field as close as possible to the actual area it now covers, changing the field boundaries as necessary.

- Sometimes you may encounter parcels of land that include items such as barns, buildings, roads, sloughs, etc. Where these items are geographically located close together and where the combined area for these items totals 3 acres or more, group them together inside one parcel and show the total acreage for all of these items inside that one parcel.

- Comment on the back of the enlarged aerial photo if you encounter any unusual circumstances, for example, poor growth in a rapeseed field, volunteer rapeseed crop in field sown with something else, flooded fields, etc. Refer to the SCR number in all cases.

- Identify separately any non-agricultural holdings in the segment, using the same definition as previously discussed in the NFS procedures, that is, land that is devoted to other than agricultural activities for example, mining, residential, churches, golf courses, etc.

- After you complete the Remote Sensing procedures, complete the remaining NFS procedures.

REMOTE SENSING PROCEDURES

SPECIAL PROCEDURES

A) Refusals and No Contacts

If you encounter any land inside a segment which belongs to a respondent who refuses to participate in the survey or a respondent whom you are unable to contact, mark the boundaries of their holdings on the enlarged aerial photo. Place the initials "NI", for Non-interview, inside the boundaries.

B) Cross Reference Situations

The NFS procedures state that, for a cross reference in a segment to a farm in another segment, you are not required to complete a full NFS segment questionnaire the second time the farmer is contacted. In the case of land operated by a specified farmer found in a segment, only the specified farm questionnaire is completed. In the case of land operated by a list farmer found in a segment, the list questionnaire is completed and the cross-reference information on the segment questionnaire.

If any of these cross-reference situations occur in one of your remote sensing segments, you are still required to complete the Remote Sensing procedures. IT IS VERY IMPORTANT THAT ALL ENLARGED AERIAL PHOTOS ARE COMPLETED AS FULLY AND AS ACCURATELY AS POSSIBLE.

C) Segment Farms on the Segment Screening List

If a segment farm in a remote sensing segment is on the Segment Screening List, you will follow regular enumeration procedures and complete the cover page, page 2, question 1 on page 3 and any other applicable questions. You must also complete the enlarged aerial photo and map the fields and crop acreages for the holding.

REMOTE SENSING PROCEDURES

FINAL CHECK

Since the material you provide is the only source of data available to train the computer to map crops from the satellite data, and since the acreages are critical to the provision of accurate numbers from the satellite, one final verification is necessary.

The final verification consists of two phases:

a) Verify the relative size of fields. Do the sizes of the fields and acreages appear consistent? Note any inconsistencies on the back of the questionnaire (Refer to SCR numbers in all cases).

b) Check that all fields have been identified on the enlarged aerial photo for all your segments.

REMOTE SENSING PROCEDURES

INFORMATION ON REMOTE SENSING

The use of Remote Sensing methods would enable Statistics Canada to produce accurate, timely data on crop acreages that would be very cost effective and result in reduced response burden for the respondent.

If any of your respondents have any questions on remote sensing or wish to get further information, instruct them to contact CCRS directly at the following address:

Canada Centre For Remote Sensing
717 Belfast Road
Ottawa, Ontario
K1A OY7

Attention: User Assistance and Marketing Unit

OR

Manitoba Centre for Remote Sensing
1007 Century Street
Winnipeg, Manitoba
R3H OW4

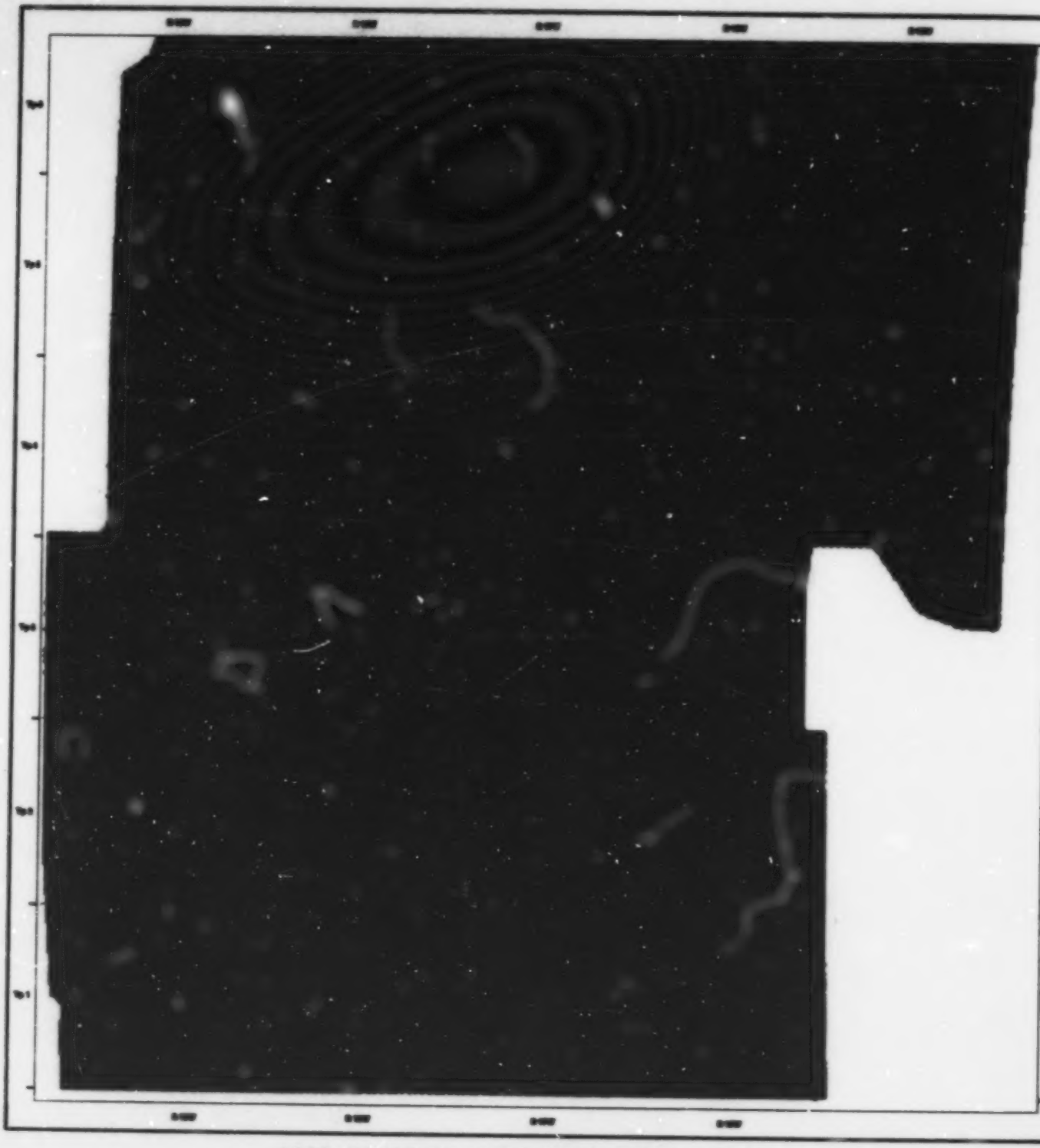
Tel: (204) 633-9543 Ext: 136

APPENDIX II

Thematic Map of Image TC1 Showing Filtered and Unfiltered Results

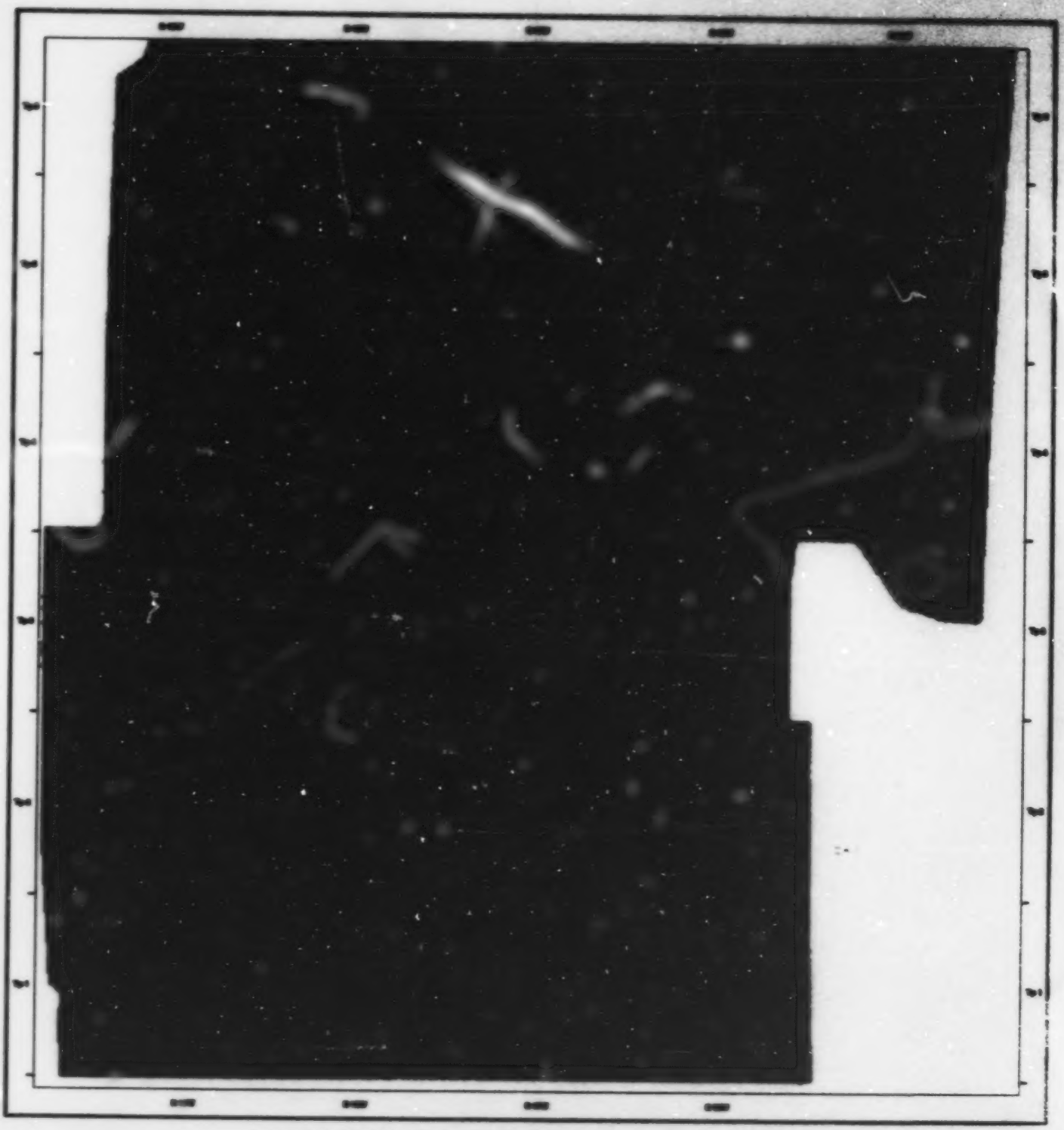
The multipurpose, multiframe NFS survey design which used total cattle and total cropland as key variables for stratification was not an optimal design for remote sensing. NFS segment locations did not always ensure adequate sampling for each major soil zone on a Landsat scene. A survey design that considered the juxtaposition of Landsat frame boundaries in relation to soil factors would result in better placement of segments.

The project has demonstrated that the infrastructure needed to handle both technical and organizational aspects of domestic crop reporting using remote sensing techniques now exists in Manitoba for regional crop reporting. Recommendations are made for expanding this program to a provincial scale.



1983 MANITOBA CROP MONITORING PROJECT

UNFILTERED



1983 MANITOBA CROP MONITORING PROJECT

FILTERED

APPENDIX III

Pixel C(by Segment for Both Filtered and
Unfilter Results for Images SWQ, TC1 and RIM.

Table 1. Pixel count for canola, summerfallow and cereal grains on filtered and unfiltered classification results for image SWQ.

SEGMENT NUMBER	FILTERED			UNFILTERED		
	RAPESEED	SUMMERFALLOW	CEREALS	RAPESEED	SUMMERFALLOW	CEREALS
5	12	77	1226	9	77	1150
6	55	2	1153	37	4	1055
7	0	28	697	0	33	704
8	0	72	1059	2	67	1055
9	20	87	458	15	82	501
10	0	302	15	0	311	48
11	0	54	152	0	55	274
12	151	147	645	132	138	652
13	0	115	427	0	117	465
14	0	364	442	0	353	451
15	44	21	852	39	23	834
16	10	251	269	7	244	305
17	0	87	212	0	88	261
21	0	73	58	0	78	151
22	76	31	429	57	31	523
23	0	202	300	0	199	330
24	110	197	402	100	186	516
25	0	78	135	0	74	168
26	0	170	707	0	159	705
27	0	277	777	2	251	755
28	0	269	10	0	275	86
29	0	58	725	0	54	717
30	0	50	521	0	56	571
31	0	402	282	0	385	330
32	0	367	390	0	345	460
33	0	248	325	2	239	318
34	270	66	983	244	59	949
35	0	32	684	0	29	686
36	44	115	955	44	102	912
37	5	201	1043	0	192	1022
38	72	98	434	67	91	504
39	0	6	0	0	7	0
40	0	4	0	0	4	0
41	0	9	0	0	9	0
42	0	13	0	0	11	0
43	3	79	202	5	71	203
44	306	2	866	288	2	876
45	86	65	1034	67	63	954
46	0	6	487	0	6	590
52	35	95	981	31	91	959
53	32	148	1049	22	142	1009
54	129	65	970	119	58	962
55	3	128	967	2	118	926
56	45	151	427	33	152	424
57	21	192	216	20	188	287
District 1 Sub Total	47477	90922	403134	41689	88684	401053
District 2 Sub Total	41731	100704	525666	36222	96739	532831
Subarea Total	89208	191626	928800	77911	185423	933884

Table 2. Pixel count for canola, summerfallow, and cereal grain on filtered and unfiltered classification results for image RIM. Note that image RIM was resampled to 50 m x 50 m pixels.

SEGMENT NUMBER	FILTERED			UNFILTERED		
	CANOLA RAPESEED	SUMMERFALLOW SAND	CEREAL GRAINS	CANOLA RAPESEED	SUMMERFALLOW SAND	CEREAL GRAINS
41	0	62	0	0	61	0
42	0	85	0	0	81	0
43	0	155	435	1	144	453
47	46	45	1100	37	55	1053
48	0	583	1135	1	517	1126
49	432	16	1685	404	19	1567
50	3	125	1027	7	98	1122
51	177	0	1031	163	2	953
52	131	186	1316	119	171	1324
53	131	971	161	133	929	276
54	0	250	1431	0	258	1373
55	63	205	1044	53	188	1183
58	91	156	1216	86	135	1255
59	0	3	1290	0	10	1364
60	266	363	557	232	326	623
61	24	55	1271	20	43	1295
Subarea Total	18619	20830	171145	16905	19406	168164

Table 3. Pixel count for canola, summerfallow and cereal grains on filtered and unfiltered classification results for image TCI.

SEGMENT NUMBER	FILTERED			UNFILTERED		
	CANOLA	SUMMERFALLOW	CEREAL GRAINS	CANOLA	SUMMERFALLOW	CEREAL GRAINS
1	122	18	593	107	14	574
2	69	3	838	85	11	720
3	31	2	747	32	4	695
4	113	2	626	105	2	568
5	18	82	654	38	69	618
15	87	29	1144	95	28	1058
18	120	62	558	117	57	535
19	0	164	825	11	153	780
20	0	95	549	2	88	468
34	367	62	411	347	54	416
36	114	123	885	112	122	823
37	8	276	1008	8	273	898
Subarea Total	29027	23379	223685	27971	21804	207046

APPENDIX IV

Detailed Account of Pixel Count by Segment for Image SWQ

Segment Analysis for Rapeseed on Image SWQ

NOTE: Discrepancies between the pixel counts in the Appendix III written analysis and those reported in Table 1 for segments 12, 15, 16 and 52 are due to removal of some portions within a segment because of refusals or non-contact situations. The numbers reported in table 1 were used for regression analysis.

The following section provides a segment by segment account of supervised classification results for canola on image SWQ. Results of both filtered and unfiltered classification are presented. Table 1 shows the number of pixels classified as canola on a segment basis for both filtered and unfiltered results.

Segment 5 Survey acreage= 0 Unfiltered= 9 Filtered= 12

The 9 pixels classified as canola form a continuous block within a large wheat field. In the vicinity of this group of canola pixels there are numerous unclassified pixels. After the filter has been applied, the contiguous block grew to 11 pixels. The other pixel in the filtered image resulted from spill over from an adjacent canola field. Depending on the positioning of the segment boundary this pixel could have been removed.

Segment 6 Survey acreage= 120 Unfiltered= 37 Filtered= 55

The 37 pixels classified as canola are positionally correct according to the segment data; however, their distribution suggests that the reported acreage may be high. The rationale for this is that the pixels are tightly packed with very few outliers. One would expect that if the entire field had been seeded to canola even a marginal canola crop would show the 37 pixels scattered over the entire 120 acres. After filtering the number of canola pixels increased to 55: however, there still remained at least 38 unclassified pixels on the 120 acre field.

Segment 7 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 8 Survey Acreage= 0 Unfiltered= 2 Filtered= 0

The two pixels classified as canola show up in a barley field but were removed by the filter.

Segment 9 Survey Acreage= 60 Unfiltered= 15 Filtered= 20

The 15 canola pixels were positionally correct according to survey data. The number of canola pixels was increased to 20 after filtering; however, there still remained 10 unclassified pixels in the vicinity of the canola field.

Segment 10 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 11 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 12 Survey Acreage= 120 Unfiltered= 184 Filtered= 210

There appears to be a significant discrepancy between reported canola acreages and the computer classified results. Classification results indicate the existance of 3 separate canola fields with each of these fields labelled as rye on the survey data. One field reported to possess 120 acres of rye contained 81 canola pixels, another 70 acre rye field contained 51 canola pixels and a third rye field of 110 acres contained 52 canola pixels. After filtering the 120 acre field increased to 90 canola pixels, the 70 acre field contained 55 canola pixels and the 110 acre field, 4 more canola pixels located under the segment boundary could be included which would bring the total pixel count for this field to 94. It was also noted that 26 pixels in the vicinity of the 110 acre field remained unclassified even after filtering.

Segment 13 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 14 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 15 Survey Acreage= 40 Unfiltered= 22 Filtered= 44

Estimates of canola for this segment were affected by haze on the satellite data. Of the 22 canola pixels, 17 were positionally correct and the other 5 were situated in a barley field that was covered by haze. With the filtered results, 26 pixels were positionally correct but the number of canola pixels in the haze area also increased to 18.

Segment 16 Survey Acreage= 260 Unfiltered= 184 Filtered= 220

A 120 acre canola field reported on the segment was not represented by any pixels classified as canola. A second, 260 acre canola field, shown on the segment data was represented by 177 canola pixels. The remaining 7 canola pixels showed up on the border between a buckwheat and wheat field. After filtering the total number of canola classified pixels for the segment increased to 220. Of these 211 were positionally correct and 9 were incorrect.

Segment 17 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 21 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 22 Survey Acreage= 90 Unfiltered= 57 Filtered= 76

All 57 canola pixels were positionally correct according to the segment data. After filtering the number of canola pixels increased to 76.

Segment 23 Survey Acreage= [illegible] Unfiltered= 0 Filtered= 0

Segment 24 Survey Acreage= 75 Unfiltered= 100 Filtered= 110

The 75 acre canola field reported on the survey was correctly classified and was represented by 56 pixels. In addition, 44 pixels were classified as canola on field 005 which was reported as a wheat field. The distribution of these 44 pixels suggests that perhaps a narrow strip of canola was grown in the field that was reported as being entirely wheat. After filtering the total number of canola classified pixels increased to 110 of which 57 were located in the position of the 75 acre field. The remaining 63 pixels were located in the wheat field. It is interesting to note that this wheat field contained approximately 75 unclassified pixels even after filtering.

Segment 25 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 26 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 27 Survey Acreage= 0 Unfiltered= 2 Filtered= 0

The 2 canola pixels appeared in a wheat field but were removed by the filter.

Segment 28 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 29 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 30 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 31 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 32 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 33 Survey Acreage= 0 Unfiltered= 2 Filtered= 0

The 2 canola classified pixels appeared in a barley field but were removed by the filter.

Segment 34 Survey Acreage= 313 Unfiltered= 244 Filtered= 270

The two large canola fields in this segment were both correctly classified. The 153 acre field was represented by 118 pixels and the 160 acre field was indicated by 126 pixels. After filtering, both fields showed an increase pixel count to 135 pixels each. There were however, 7 more pixels located under the segment border for the 153 acre field, bringing its total to 142 pixels and 3 more segment border pixels for the 160 acre field, bringing its total to 138 pixels. In both instances, pixels taht were missed in the count because they fell under the segment border, likely belonged to the two canola fields contained in this segment because there were no adjacent canola fields outside the segment.

Segment 35 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 36 Survey Acreage= 60 Unfiltered= 44 Filtered= 44

Survey data for this segment indicated two 30 acre fields separated by a natural drainage ditch. One of the fields was correctly identified as canola and was represented by 17 pixels; however, the other field represented by 23 pixels did not appear to be positionally correct. The 4 other pixels classified as canola appeared in a sunflower field.

After filtering the total pixel count for the segment remained at 44. The field originally represented by 17 pixels gained 2 more pixels and the field of 23 pixels gained 1 more pixel. One more pixel entered the segment as a result of spill over from an adjacent canola field. All those canola pixels that appeared in the sunflower field were removed by the filter.

Segment 37 Survey Acreage= 0 Unfiltered= 0 Filtered= 5

The 5 pixels classified as canola after filtering were a result of spill-over from an adjacent canola field. Repositioning of the segment border could perhaps reduce this number.

Segment 38 Survey Acreage= 85 Unfiltered= 67 Filtered= 72

The 67 pixels classified as canola were positionally correct according to segment data.

After filtering the number of canola pixels increased to 72. An additional 3 more pixels 'missed' as a result of being located under the segment border, could be added to bring the total pixel count for this field to 75.

Segment 39 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

This segment covered the southern district of the Town of Souris. Rationale for this segment selection for remote sensing purposes, is unknown.

Segment 40 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

This segment covered the northern district of the Town of Souris.

Segment 41 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

This segment covered a populated portion of the City of Brandon.

Segment 42 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

This segment covered a populated portion of the City of Brandon.

Segment 43 Survey Acreage= 0 Unfiltered= 5 Filtered= 3

Four of the canola pixels appeared in a field marked on the segment data as N/A. The other pixel was "out standing" in a field of alfalfa. After filtering only 3 of the pixels in the N/A field remained.

Segment 44 Survey Acreage= 350 Unfiltered= 288 Filtered= 306

All 288 canola pixels were positionally correct according to segment data. I suspect that the actual acreage of canola was less that the reported 350 acres as a second yard site and other land along the northern fringe of the field did not appear to have been considered. After filtering the total number of canola pixels increased to 306.

Segment 45 Survey Acreage= 103 Unfiltered= 67 Filtered= 86

The exact location of the canola fields on the segment data was difficult to establish; however, 64 pixels appeared to be positionally correct. The remaining 3 pixels appear to have been misclassified as barley and flax. After filtering, these 3 pixels were removed and the total pixel count for canola increased to 86 pixels.

Segment 46 Survey Acreage= 0 Unfiltered= 0 Filtered= 0

Segment 52 Survey Acreage= 90 Unfiltered= 66 Filtered= 89

The relationship between pixel count and reported acreage in this segment is misleading. The 90 acre canola field in the northwest corner of the segment has been classified as canola and is comprised of 32 pixels. The general shape of the field has been accurately portrayed; however, within the field the pixel distribution is spotty. Numerous potholes on the photograph lead one to believe that the reported acreage for this field might be overstated. After filtering the number of canola pixels for this field increased to 35.

An 8[illegible] acre field marked on the segment data as <u>not rapeseed</u> (004) was classified as canola and received 44 pixels. After filtering the number of canola pixels increased to 54.

Segment 53 Survey Acreage= 140 Unfiltered= 22 Filtered= 32

All 22 Canola pixels from the unfiltered classification were positionally correct. After filtering the canola pixel count increased to 32.

There were 15 unclassified pixels around the canola field.

Segment 54 Survey Acreage= 152 Unfiltered= 119 Filtered= 129

There were 119 canola pixels which were positionally correct; however, the classification appears spotty. In addition there were 5 pixels classified as canola that were found under the segment border. If these 5 pixels were included, the total pixel count would be 124.

After filtering the total canola pixel count increased to 129. If the 7 canola pixels under the segment border were included, the total pixel count would be 136.

Segment 55 Survey Acreage= 0 Unfiltered= 2 Filtered= 3

The 2 canola pixels in this segment were the result of spill-over from an adjacent canola field. After filtering this number increased to 3 pixels. These pixels could likely be removed by redrawing segment boundaries.

Segment 56 Survey Acreage= 50 Unfiltered= 33 Filtered= 45

The 33 canola pixels were positionally correct according to segment data. The classification appears somewhat spotty but after filtering the field is well marked and represented by 45 canola pixels.

Segment 57 Survey Acreage= 69 Unfiltered= 20 Filtered= 21

Three canola fields were reported in the survey data; however, only two fields were classified as canola on the satellite data. The field reported as 29 acres was represented by 11 pixels and the remaining 9 pixels were located on a 20 acre canola field. One 20 acre canola field was not represented by any canola pixels. After filtering the field originally represented by 11 pixels increased to 13 pixels and the other field of 9 pixels decreased by one pixel.

SUMMARY

Although several discrepancies between segment and satellite data exist, the overall classification results for canola look good. There were only two instances where entire canola fields reported in the survey were not classified on the satellite data.

A 120 acre canola field and a 20 acre canola field reported for segments 16 and 57 respectively were not represented by any canola classified pixels.

There were 5 cases where large numbers of canola pixels appeared in fields that were reported in the survey as being something other than canola. In segments 5, 16 and 24 canola pixels showed up in wheat fields. In segment 12 canola pixels appeared in 3 different rye fields. The latter error was likely due to incorrect labelling of the field on the survey photograph. Finally in segment 52, a large number of pixels appeared in a field marked simply as "not rapeseed".

Other trends in the classification and filtering results were noted. For example, in segment numbers 6, 9, 12, 24 and 53 there were numerous unclassified pixels associated with the canola fields. This indicated that perhaps the training statistics for canola could be expanded. In segment numbers 12, 34, 38, 54 and 55 the total pixel count could be altered if border pixels were included in the pixel count.

In several instances where only a few pixels had been misclassified (i.e. segments 8, 27, 33, 36 and 45) the filtering process often removed these anomalies. After filtering the number of canola pixels in segments 5, 36, 37 and 55 increased as a result of adjacent canola fields "growing" in size and spilling over into the segment. Classification results for segment 15 were seriously affected by haze. Segment numbers 39, 40, 41 and 42 could not be used for crop signature development because these areas contained urban centres and possessed no crops.